Towards establishment of a centralized spider traits database

Lowe, Elizabeth C.

2020-09

Lowe , E C , Wolff , J O , Aceves-Aparicio , A , Birkhofer , K , Branco , V V , Cardoso , P , Chichorro , F , Fukushima , C S , Goncalves-Sousa , T , Haddad , C , Isaia , M , Krehenwinkel , H , Audisio , T L , Macias Hernandez , N , Malumbres-Olarte , J , Mammola , S , McLean , D J , Michalko , R , Nentwig , W , Pekar , S , Petillon , J , Privet , K , Scott , C , Uhl , G , Urbano Tenorio , F , Wong , B H & Herbestein , M E 2020 , ' Towards establishment of a centralized spider traits database ' , Journal of Arachnology , vol. 48 , no. 2 , pp. 103-109 . https://doi.org/10.1636/0161-8202-48.2.103

http://hdl.handle.net/10138/332724
https://doi.org/10.1636/0161-8202-48.2.103

unspecified
acceptedVersion

Towards establishment of a centralized spider traits database

Elizabeth Lowe*[1], Jonas O. Wolff*[1], Alfonso Aceves-Aparicio[1], Klaus Birkhofer[2], Vasco Veiga Branco [3], Pedro Cardoso[3,4], Filipe Chichorro[3], Caroline Sayuri Fukushima[3], Thiago Gonçalves-Souza[5], Charles R. Haddad[6], Marco Isaia[7], Henrik Krehenwinkel[8], Tracy Lynn Audisio[9], Nuria Macías-Hernández[3,18], Jagoba Malumbres-Olarte[3,4], Stefano Mammola[3,10], Donald James McLean[1], Radek Michalko[11], Wolfgang Nentwig[12], Stano Pekár[13], Julien Pétillon[14], Kaïna Privet[14], Catherine Scott[15], Gabriele Uhl[16], Fernando Urbano-Tenorio[3], Boon Hui Wong[17], and Marie E. Herberstein[1]

* Joint first author and corresponding authors: lizzy.lowe@mq.edu.au, Jonas.wolff@mq.edu.au

1. Department of Biological Sciences, Macquarie University, NSW 2109, Australia

2. Department of Ecology, Brandenburg University of Technology Cottbus-Senftenberg, 03046 Cottbus, Germany

3. Laboratory for Integrative Biodiversity Research (LIBRe), Finnish Museum of Natural History (Luomus), University of Helsinki, P.O.Box 17 (Pohjoinen Rautatiekatu 13), 00014 Helsinki, Finland.

4. cE3c – Centre for Ecology, Evolution and Environmental Changes / Azorean Biodiversity Group and Universidade dos Açores – Faculty of Agriculture and Environment, Rua Capitão João d'Ávila, São Pedro, 9700-042 Angra do Heroísmo, Terceira, Açores, Portugal.

5. Ecological Synthesis and Biodiversity Conservation Lab, Department of Biology, Federal Rural University of Pernambuco, Brazil

6. Department of Zoology & Entomology, University of the Free State, P.O. Box 339, Bloemfontein, 9300, South Africa

7. Laboratory of Terrestrial Ecosystems, Department of Life Science and Systems Biology, University of Turin, Italy

8. Department of Biogeography, University of Trier, 54296-Trier, Germany

9. Evolutionary Genomics Unit, Okinawa Institute of Science and Technology Graduate University, Onna-son, Okinawa, Japan

10. Molecular Ecology Group (MEG), Water Research Institute (IRSA), National Research Council of Italy (CNR), 28922 Verbania Pallanza, Italy

11. Department of Forest Ecology, Faculty of Forestry and Wood Technology, Mendel University in Brno, Zemědělská 3, 613 00 Brno, Czech Republic

12. Institute of Ecology and Evolution, University of Bern, CH-3012 Bern, Switzerland

13. Department of Botany and Zoology, Faculty of Science, Masaryk University, 611 37, Brno, Czech Republic

14. Université de Rennes, CNRS, Ecobio (Ecosystèmes, biodiversité, évolution) – UMR 6553, F-35000 Rennes, France

15. Department of Biology, Acadia University, 15 University Ave, Wolfville, Nova Scotia, Canada B4P 2R6

16. Zoologisches Institut und Museum Allgemeine und Systematische Zoologie, Greifswald, Germany

17. Behavioural Ecology and Sociobiology Lab, Department of Biological Sciences, National University of Singapore, Singapore.

18. Island Ecology and Evolution Research Group, IPNA-CSIC, Tenerife, Canary Islands, Spain.

Abstract. A main goal of ecological and evolutionary biology is understanding and predicting interactions between populations and both abiotic and biotic environments, the spatial and temporal variation of these interactions, and the effects on population dynamics and performance. Trait-based approaches can help to model these interactions and generate a comprehensive understanding of ecosystem functioning. A central tool is the collation of databases that include species trait information. Such centralized databases have been set up for a number of organismal groups but is lacking for one of the most important groups of predators in terrestrial ecosystems – spiders. Here we promote the collation of an open spider traits database, integrated into the global Open Traits Network. We explore the current collation of spider data and cover the logistics of setting up a global database, including which traits to include, the source of data, how to input data, database governance, geographic cover, accessibility, quality control and how to make the database sustainable long-term. Finally, we explore the scope of research questions that could be investigated using a global spider traits database.

Keywords: phenotypic traits, functional diversity, functional ecology, ecosystem functioning, evolutionary ecology, comparative analysis

We are living in an age of large-scale assembly of digital information. The generation of large databases across scientific disciplines is now enabling the modeling of natural processes on an unprecedented scale. The development of the world wide web over the last three decades has aided the centralization of collated data and has enabled the storing and sharing of large amounts of data. Biology has truly arrived in the information age with the large-scale digitization of information and its central storage in online databases (e.g. GenBank, BOLD, Dryad). These databases have become the core of many recent research advances.

In ecology and evolutionary biology, the accessibility of large-scale genetic information reaches its true potential when linked with phenotypic and faunistic data. For instance, comparative studies of organismal traits among species or across space, time or habitat types, can reveal general patterns that help to predict changes in biodiversity and ecosystem functioning in a rapidly changing world (Krause et al. 2014; Madin et al. 2016b; Wright et al. 2017; Dudley et al. 2019; Mammola et al. 2019). However, the development of open databases for phenotypic data is lagging far behind. This not only limits progress in the life sciences, but also leads to inefficient duplication and impairs reproducibility of research. Major reasons for the slow and reluctant establishment of trait databases are the perceived difficulty to unify data and to standardize reporting methods. In contrast, for molecular data these challenges have been overcome, facilitating the rapid growth of molecular databases. Furthermore, the inclusion of new molecular data into such central databases is highly promoted. For example, final acceptance of a publication in many scientific journals is contingent on the submission of genetic data to a genetic database (e.g. GenBank).

Compared to genetic data, trait data are much harder to standardize across taxonomic groups due to different terminologies and methods used for their measurement and recording. Therefore, existing trait database projects are taxon-specific and often constrained to specific traits or geographic regions (e.g. (Kattge et al. 2011; Homburg et al. 2014; Brun et al. 2016; Madin et al. 2016a; Oliveira et al. 2017; Parr et al. 2017)). There are good reasons to continue with the assembly of taxon-specific trait databases as this approach offers the flexibility required to tailor the database to the need of the focus field (e.g. entomology, herpetology, botany, microbiology). The data deposited in different taxon-specific databases could be synthetized into a central node by the identification of common or equivalent traits that can be easily translated into cross-taxon standards (Gallagher et al. 2020). Examples of such traits already in use in cross-taxon studies are body size, trophic niche, habitat preferences, and metabolic rates.

Hence, the building of taxon-specific trait databases can bring enormous benefits both for the specific field compiling the data (e.g. arachnology) as well as for cross-taxon analyses. Here, we propose to set

up a community-based spider trait database (Arachnida: Araneae) and invite submissions of data from the arachnological community.

Spiders are a mega-diverse order of animals, with about 48,400 species currently identified (World Spider Catalog 2020), and one of the most dominant groups of invertebrate predators in terrestrial ecosystems (Turnbull 1973). They thus play a key role in ecosystem function worldwide, such as increasing crop performance by suppressing pest density (Michalko et al. 2019a). Spiders are premier subjects to study fundamental biological mechanisms, such as sexual selection (Herberstein et al. 2013), the evolution of sociality (Lubin & Bilde 2007) and extended phenotypes (Blackledge et al. 2009), predator-prey interactions (Pekár & Toft 2015), community assembly (Birkhofer et al. 2017), and migration (Gillespie et al. 2012). They are also in the focus of the applied sciences, such as agro-ecology (pest control) (Uetz et al. 1999; Birkhofer et al. 2016; Michalko et al. 2019b), material science (spider silk) (Wolff et al. 2017) and medical science (venom) (King & Hardy 2013).

GENERAL DATABASE BENEFITS

Trait databases are particularly useful for comparative studies and studies on biodiversity and the function of ecological communities. Traditionally, such research relied on the primary collation of comparative data by laborious measurements in a large number of species and specimens or time-consuming gathering from print-literature. The widespread lack of primary or even derived data has not only contributed to the reproducibility crisis of science (Allison et al. 2018) but also led to an unnecessary redundancy in data acquisition, wasting time and money. Only recently has it become a common good scientific practice to make the raw data that underlie a study available to the scientific community. While this enhances transparency, it does not necessarily solve the problem of work redundancy as such datasets are largely hidden in appendices, electronic supplemental material or on file servers instead of a central repository. The building of a commonly accepted trait database is a solution.

Having open access to a database of relevant trait data not only boosts comparative and ecological studies, but also allows researchers to undertake quick exploratory analyses to generate preliminary data that may be useful for project planning and funding justification in grant applications. Well populated databases also make it possible to detect knowledge gaps or geographical biases, which can be used to identify future areas of research.

Trait databases thus fill an important gap in the digitization and accessibility of biological knowledge and can become a powerful tool when linked with existing taxonomic (World Spider Catalog 2020), semantic (Spider Anatomy Ontology—SPD (Ramírez & Michalik 2019)), molecular (BOLD (Ratnasingham & Hebert 2007), Genbank (Benson et al. 2012), ArachnoServer (Pineda et al. 2018)) and faunistic (e.g. Atlas of the European Arachnids; Atlas of Living Australia; Araneae Spiders of Europe (Nentwig et al. 2020); Global Biodiversity Information Facility—GBIF) resources (See Table. 1 for current databases that include spider traits). Finally, a trait database can also serve as a central data repository to fulfil the requirements of data accessibility.

BENEFITS OF A SPIDER TRAITS DATABASE

Setting up a global trait database for spiders will elevate arachnology to the next level of biodiversity digitization: having open and better access to data will bring arachnology together and create stronger collaborative networks able to address a wide range of questions (as explored below). It will make global or large-scale studies possible within a fraction of the time and improve the statistical power of many analyses. A centralized database will also help to identify species' vulnerability and guide conservation decisions (Gonçalves-Souza et al. 2015). It can also help inform taxonomic decisions since the database acts as a source of information about variability (e.g. geographic and morphological).

Beyond arachnology, a spider trait database will facilitate the inclusion of spiders in global multi-taxa studies and biomonitoring programs. The topical research area of biodiversity-ecosystem functioning relationships often utilizes metrics of functional diversity that are based on traits in cross taxon studies. The collation of global data will provide better information for outreach purposes. For

example, the database could provide information about local diversity of spiders, size comparisons or peculiar species that are accessible to local residents. A comprehensive understanding of the functional and ecological diversity of spiders, if adequately published for the general audience, can help to improve the public perception of spiders.

As existing trait database projects are largely family-focused and no such database yet exists for any arthropod order, the arachnological community could lead a frontier project for higher-level animal trait databasing. We think that arachnologists are in an excellent position to swiftly establish a spider trait database due to its outstanding achievement of digitization of all taxonomical literature (Nentwig et al. 2015) and the existence of a digital taxonomic standard, the World Spider Catalog (WSC) (2020), which makes possible the large-scale extraction of basic traits, such as area of distribution or body size (e.g. through text mining approaches), and their automatically updated taxonomic assignment (through a unique LSID identifier).

LOGISTICS OF A TRAIT DATABASE

Some of the logistics of setting up a trait database require consensus on: trait definitions, the source of data, how to input data, database governance, geographic cover, accessibility, quality control, financial needs, and how to achieve the database's long term sustainability. Below we propose principles that, in our view, will make a useful framework for a spider trait database.

(1) Definition of traits. - One of the fundamental problems with collating traits in particular is the confusion and debate around the term "traits", and in particular "functional traits" (Violle et al. 2007; Pey et al. 2014; Moretti et al. 2017; Brousseau et al. 2018; Wong et al. 2019). At their most basic form, all observable and quantifiable characters are traits. While some traits directly imply an ecological function (e.g. 'habitat preference' (Rusch et al. 2015)), for basic phenotypic data, such as body shape, the link to function is often not experimentally established and remains hypothetical. Furthermore, to be useful for a wide range of topics beyond functional ecology, a database should be flexible enough

to accommodate a diversity of data types. For a spider trait database, we propose that the term 'trait' should thus be applied in a broad sense and include all observable and quantifiable characters on the level of individuals, populations or species, irrespective of our current assumptions on their functional role.

For a highly versatile database, we propose the inclusion of the following trait categories: (a) Ecological (e.g. (micro-)habitat preference, hunting mode), (b) Behavioral (e.g. stratum utilization, maternal care), (c) Morphological (e.g. body length, leg length), (d) Physiological (e.g. reproductive rate, resting metabolic rate, thermal limits) and (e) Biomechanical (e.g. running speed, silk strength). In the future the database may also be expanded to include more categories.

(2) Data types. - Another challenge is that the nature of traits can be diverse: traits can be quantified and described using data that are descriptive (e.g. Latin name of a prey item), categorical (e.g. web-builder, active hunter), ordinal (e.g., small, medium, large), interval (e.g. 1–3, 4–6), ratio (e.g. 10:100; 20:100), discrete (e.g. 1,2,3) or continuous (e.g. 1, 3.5, 7.2). There are many examples where qualitative information is sufficient in order to capture the traits, but problems arise when researchers apply different terminologies, definitions and categorizations for observations or characteristics. While common standards must be identified for each trait to ensure comparability (e.g. fixing units for numerical traits), the versatility of trait input can be maintained by a liberal trait definition with the accurate method (e.g. recording scheme) defined in the meta data, which permits the filtering for comparable data at the analysis stage (Jones et al. 2006). Another solution is to handle differently recorded trait data under separately defined traits.

The use of non-equivalent data types for some traits presents further difficulties. For instance, different authors report phenology, habitat preferences or prey type as binary or in frequencies. Furthermore, the state of some categorical traits might be ambiguous or context dependent. For instance, in many orb-web spiders (Araneidae), the adult males abandon webs, while some long-jawed orb-web spiders (Tetragnathidae) can be both web-building or free hunting. This problem can be

solved by recording any categorical trait as binary with the option to provide a frequency for each character state and a record of the sample size. Some traits are covarying, e.g. body length and leg length. This can be accounted for by linking trait records that are derived from the same object (i.e. specimen, individual or population).

(3) Intraspecific variation. - The impact of intraspecific variation on comparative approaches has been widely demonstrated (Garamszegi & Møller 2010) and has also been recognized in community ecology (Bolnick et al. 2011). Some spider species, such as *Argiope bruennichi* (Scopoli, 1772) (Araneidae) have a wide distribution range and show an enormous variation in phenology, morphology and behavior depending on the geographical location (Krehenwinkel & Tautz 2013; Wolz et al. 2020). In such instances, a single trait record may not be representative for a species. The trait database can provide the opportunity to accommodate multiple records per species and attach relevant information about the record context (such as locality, date and developmental stage) in the metadata.

(4) Taxonomic integrity. - A serious challenge for the maintenance of multi-species datasets is taxonomic changes. Fortunately, with the World Spider Catalog (2020), arachnology has a reliable and maintained source for changing taxonomic information. A key is the assignment of a unique identifier (the Life Species Identifier (LSID)), which is constant and can be used as a link between the spider trait database and the World Spider Catalog to automatically update the taxonomic information in the database. To make any changes traceable, we propose that species names originally used in uploaded datasets must be kept in the metadata (published species name from the trait sources vs. valid species name from the WSC).

(5) Database governance and data upload. - The spider trait database will be a community-driven project, created and maintained by a team of 5-10 administrators and a webmaster. To upload data the contributor must use a template that provides the structure for how the data and attached

metadata are reported. An annotated draft template can be found in the electronic supplemental material of this article. At this stage, any trait data can be sent to the corresponding authors of this paper using this template for its inclusion into a first version of the spider trait database. The database will be hosted by Masaryk University, Czech Republic, and curated by Stano Pekár's research team and collaborators.

(6) Quality control. - Databases require some degree of curation in order to maintain their integrity. While some of the data checking can be automated (i.e. if all mandatory fields contain information and data types for traits are correct), we think that the content of submissions must be formally cross-checked and approved by an expert before it is published in the database. This may include a quick check, if the uploader is a known expert, if the data are of relevance for the database, if the data are accompanied by a citable publication that described the methods of data acquisition and that there are no obvious errors (e.g. false units and digits). The formation of an editorial board with experts for specific types of traits (e.g. behavior, morphology, habitat) will help to facilitate the described curation task.

(7) Accessibility. - Following the principles of FAIRness and open science (Wilkinson et al. 2016) and facilitating its embedding in the Open Traits Network (Gallagher et al. 2020), we promote open accessibility of the spider trait database. This can be enabled by the use of the creative commons license CC-BY 4.0, which means that anyone will be able to download and use the data for their own research with the appropriate citation of the database.

(8) Author attribution. - Authorship models and citation rules for database usage are matter of dispute at trait database meetings. The reason is that the processing and integration of data from thousands of sources makes it hard to define what makes a significant contribution that merits authorship and who must be named and cited according to the ethical norms of science and the CC-BY license

(Gallagher et al. 2020). Here it can be useful to distinguish between the authors of the original data sources (i.e. the publication from which data were extracted to build a comparative dataset) and the authors of the data synthesis (i.e. researchers who collated the comparative data set and/or made the database accessible in a systematic manner). For practical reasons, we propose that only the latter merits authorship of the database.

Following the above standards of reproducibility and traceability, it will be a requirement to cite the original data sources (preferably with a DOI) in the database. However, among the research community there is no consensus about whether original sources of legacy data must be cited in a publication that uses data from the database, and this decision is often left to the users (Gallagher et al. 2020). On the other hand, it is commonly accepted that any outputs derived from the trait database will reference the database (the synthesis). In order to become citable, the database will be released alongside an article in a scientific journal that describes the database. For the spider trait database, it is aimed to publish its formal description in a major Open Access journal. After the release, the database will be continuously updated and its content expanded. Major changes to its content and the author team will require the occasional publication of a new article under a changed version number of the database.

ADDITIONAL CHALLENGES FOR A SPIDER TRAIT DATABASE

The first major challenge of creating a spider trait database stems from the huge diversity of spiders (48,400 described species and many yet to be discovered), largely preventing arachnologists from gathering traits for all of them. One solution to this problem is to integrate data from related species in the analysis step, in a careful and controlled manner, which is a widespread procedure in comparative analyses of invertebrates (e.g. (Madin et al. 2016b)).

Second, many ecological studies use “morphospecies”, where species identification is considered too difficult. If species names are lacking, the common standard that makes the data comparable is missing. We argue that the arbitrary definition of morphospecies cannot replace species identification

by experts, e.g. by comparison with species descriptions and type material, and we do not promote the inclusion of trait data from morphospecies into the database. However, to avoid the exclusion of data from lesser known taxa, the database will allow the use of genus levels as a common standard (i.e. usage of the genus LSID, if species identities are unknown or yet to be described). It will be encouraged to specify voucher IDs and locations, if present, in the meta data, to enhance the tracing of species identifications.

POTENTIAL RESEARCH THAT WOULD BECOME FEASIBLE WITH A SPIDER TRAITS DATABASE

Metacommunity, macroecology and biogeography. - At its most basic level, a traits database will help answer questions about how traits are geospatially distributed across taxa. Studies on how traits evolve or change at the community level in response to natural and historic change would also be facilitated. Having access to data relating to trait composition and body size will enable researchers to study the macroecological drivers of functional diversity at broad spatial and temporal scales (Cardoso et al. 2011). It will also enable the study of the relative contribution of environmental filtering, dispersal-related processes and competitive exclusion in determining spider assemblages and their functional diversity (Jiménez-Valverde et al. 2010; Carvalho & Cardoso 2014). How the mechanisms of environmental filtering influence community assembly along environmental gradients (e.g. climatic, altitudinal, latitudinal, structural) and by abiotic and biotic disturbances (e.g. floods, fire and biological invasions) can also be investigated by integrating data from the trait database with data from other online resources, such as climatic databases (Carvalho et al. 2020). The integration of vegetation biome data and climatic gradients allows for the testing of the relationship between vegetation, latitude and functional diversity of spider communities. Finally, the effects of urbanization on the expression or evolution of spider traits could be studied on a global scale (Lowe et al. 2018).

Evolution. - To date, trait-based research is a topic largely promoted within macro-ecology, and less recognized in comparative biology. However, the digitization of traits has considerable potential for the fields of comparative, evolutionary and systematic biology. By linking phylogenomic data with trait

data, fundamental evolutionary questions can be studied on a large taxonomic scale, such as the evolution of body size (Kuntner & Coddington 2019) and the extended phenotypes (Blackledge et al. 2009). It would also be interesting to examine connections between phylogenetic diversity and functional diversity and to test the 'stable species hierarchy hypothesis' in spiders, which predicts that trait variation is higher at the interspecific level than at the intraspecific level (Kazakou et al. 2014). Furthermore, testable ideas include whether diversification processes have been accompanied by expansions or shifts in trait space and how the evolution of different traits are correlated.

Ecology and ecosystem functioning. - Trait based research can help to understand how communities and ecosystems work by unraveling the interactions between species and their ecological roles. With a comprehensive spider trait database, it could be feasible to identify which traits are associated with specific habitat types, such as the ones found in caves, deserts or mountain summits. In combination with similar databases of other taxa, the traits affecting ecosystem functioning, those affected by ecosystem processes, and traits that are redundant could be identified. Also, functional traits that have a similar (or complementary) ecological function could be identified. From an applied perspective, understanding which traits contribute to high levels of pest control by spiders ("effect traits", Rusch et al. 2015) and how this service may be supported by certain management practices ("response traits" (Birkhofer et al. 2014)) is an important research field in agriculture and forestry. Questions of niche formation in a competitive setting could be addressed, for instance what are the traits that confer competitive advantages under different circumstances.

Conservation. - The proposed spider trait database can guide conservation decisions. By identifying species that exhibit a unique combination of traits, local conservation can guide efforts targeting these species (Birkhofer et al. 2017). Comparative studies can reveal traits that are associated with species' extinction risk, resilience or the recolonization of restored areas, for instance using model species' vulnerability to climate and habitat changes (Chichorro et al. 2019). Future threats to biodiversity from land-use change or intensification could become predictable, if we understood community responses from a trait perspective. Comparing both taxonomic and functional diversity responses to agricultural

intensification provides an ideal framework to consider trade-offs and synergies between different conservation goals (Birkhofer et al. 2015). To understand which species are threatened by the pet trade, the database can identify the traits, such as size or color, that make a spider attractive to the buyer.

OUTLOOK

This paper demonstrates how the arachnological community and biologists more broadly would benefit from a global spider traits database. We have discussed the challenges that the development of a comprehensive trait database brings for a megadiverse group such as spiders and have provided possible solutions for each problem. We suggest that establishing the spider traits database is feasible, but will require major support, endorsement and input by the community. Following a fruitful workshop on 10th February 2019, at the 21st International Congress of Arachnology in Christchurch, New Zealand, an action group has formed that is currently working on the implementation of the principles discussed in this article. The team, led by Stano Pekár, is currently setting up the online database, with the release of a first version planned for the end of 2020. We are requesting both data submissions (using the provided template) and participation in editing, testing and data curation. For questions, suggestions and offers please contact the corresponding authors of this article.

ACKNOWLEDGEMENTS

This paper was written as a result of a workshop at the International Congress of Arachnology in Christchurch, New Zealand 2019. We would like to thank the ICA organizing committee and the International Society of Arachnology for their support, the Macquarie University Species Spectrum Research Center and a Pentti Tuomikoski Grant from the Finnish Museum of Natural History for financial support. JOW was supported by a Discovery Early Career Researcher Award of the Australian Research Council (DE190101338).

LITERATURE CITED

Allison D.B., R.M. Shiffrin & V.J.P.o.t.N.A.o.S. Stodden 2018. Reproducibility of research: Issues and proposed remedies. *Proceedings of the National Academy of Sciences* 115:2561-2562.

Benson D.A., M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell et al. 2012. GenBank. *Nucleic acids research* 41:D36-D42.

Birkhofer K., F. Arvidsson, D. Ehlers, V. Mader, J. Bengtsson & H. Smith 2016. Organic farming affects the biological control of hemipteran pests and yields in spring barley independent of landscape complexity. *Landscape ecology* 31:567-579.

Birkhofer K., T. Diekötter, C. Meub, K. Stötzel & V. Wolters 2015. Optimizing arthropod predator conservation in permanent grasslands by considering diversity components beyond species richness. *Agriculture, Ecosystems Environmental Entomology* 211:65-72.

Birkhofer K., M.M. Gossner, T. Diekötter, C. Drees, O. Ferlian, M. Maraun et al. 2017. Land-use type and intensity differentially filter traits in above-and below-ground arthropod communities. *Journal of Animal Ecology* 86:511-520.

Birkhofer K., V. Wolters & T. Diekötter 2014. Grassy margins along organically managed cereal fields foster trait diversity and taxonomic distinctness of arthropod communities. *Insect Conservation Diversity* 7:274-287.

Blackledge T.A., N. Scharff, J.A. Coddington, T. Szüts, J.W. Wenzel, C.Y. Hayashi et al. 2009. Reconstructing web evolution and spider diversification in the molecular era. *Proceedings of the National Academy of Sciences* 106:5229-5234.

Bolnick D.I., P. Amarasekare, M.S. Araújo, R. Bürger, J.M. Levine, M. Novak et al. 2011. Why intraspecific trait variation matters in community ecology. *Trends in ecology, evolution & development* 26:183-192.

Brousseau P.M., D. Gravel & I.T. Handa 2018. On the development of a predictive functional trait approach for studying terrestrial arthropods. *Journal of Animal Ecology* 87:1209-1220.

Brun P.G., M.R. Payne & T. Kiørboe 2016. A trait database for marine copepods. *Earth System Science Data Discussions*:1-33.

Cardoso P., S. Pekár, R. Jocqué & J.A. Coddington 2011. Global patterns of guild composition and functional diversity of spiders. *PloS one* 6:e21710.

Carvalho J.C. & P. Cardoso 2014. Drivers of beta diversity in Macaronesian spiders in relation to dispersal ability. *Journal of Biogeography* 41:1859-1870.

Carvalho J.C., J. Malumbres-Olarte, M.A. Arnedo, L.C. Crespo, M. Domenech & P. Cardoso 2020. Taxonomic divergence and functional convergence in Iberian spider forest communities: insights from beta diversity partitioning. . *Journal of Biogeography* In Press

Chichorro F., A. Juslén & P. Cardoso 2019. A review of the relation between species traits and extinction risk. *bioRxiv*:408096.

Dudley A., N. Butt, T.D. Auld & R.V. Gallagher 2019. Using traits to assess threatened plant species response to climate change. *Biodiversity and Conservation* 28:1905-1919.

Gallagher R.V., D.S. Falster, B.S. Maitner, R. Salguero-Gómez, V. Vandvik, W.D. Pearse et al. 2020. Open Science principles for accelerating trait-based science across the Tree of Life. *Nature ecology & evolution* 4:294–303.

Garamszegi L.Z. & A.P. Møller 2010. Effects of sample size and intraspecific variation in phylogenetic comparative studies: a meta-analytic review. *Biological Reviews* 85:797-805.

Gillespie R.G., B.G. Baldwin, J.M. Waters, C.I. Fraser, R. Nikula & G.K. Roderick 2012. Long-distance dispersal: a framework for hypothesis testing. *Trends in ecology and evolution* 27:47-56.

Gonçalves-Souza T., A.J. Santos, G.Q. Romero & T.M. Lewinsohn 2015. Conservation along a hotspot rim: spiders in Brazilian coastal restingas. *Biodiversity and Conservation* 24:1131-1146. doi:10.1007/s10531-014-0846-8

Herberstein M.E., E. Hebets & D. Penney 2013. Why are spiders good models for research. In: Penney D (ed) Spiders in the 21st Century Siri Scientific Press, pp 230-251.

Homburg K., N. Homburg, F. Schäfer, A. Schuldt & T. Assmann 2014. Carabids. org–a dynamic online database of ground beetle species traits (Coleoptera, Carabidae). *Insect Conservation and Diversity* 7:195-205.

Jiménez-Valverde A., A. Baselga, A. Melic & N. Txasko 2010. Climate and regional beta-diversity gradients in spiders: dispersal capacity has nothing to say? *Insect Conservation and Diversity* 3:51-60.

Jones M.B., M.P. Schildhauer, O. Reichman & S. Bowers 2006. The new bioinformatics: integrating ecological data from the gene to the biosphere. *Annual Review of Ecology, Evolution and Systematics* 37:519-544.

Kattge J., S. Diaz, S. Lavorel, I.C. Prentice, P. Leadley, G. Bönisch et al. 2011. TRY–a global database of plant traits. *Global change biology* 17:2905-2935.

Kazakou E., C. Violle, C. Roumet, M.L. Navas, D. Vile, J. Kattge et al. 2014. Are trait-based species rankings consistent across data sets and spatial scales? *Journal of Vegetation Science* 25:235-247.

King G.F. & M.C. Hardy 2013. Spider-venom peptides: structure, pharmacology, and potential for control of insect pests. *Annual review of entomology* 58:475-496.

Krause S., X. Le Roux, P.A. Niklaus, P.M. Van Bodegom, J.T. Lennon, S. Bertilsson et al. 2014. Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Frontiers in microbiology* 5:251.

Krehenwinkel H. & D. Tautz 2013. Northern range expansion of E uropean populations of the wasp spider A rgiope bruennichi is associated with global warming–correlated genetic admixture and population-specific temperature adaptations. *Molecular ecology* 22:2232-2248.

Kuntner M. & J.A. Coddington 2019. Sexual Size Dimorphism: Evolution and Perils of Extreme Phenotypes in Spiders. *Annual Review of Entomology* 65

Lowe E.C., C.G. Threlfall, S.M. Wilder & D.F. Hochuli 2018. Environmental drivers of spider community composition at multiple scales along an urban gradient. *Biodiversity and Conservation* 27:829-852. doi:10.1007/s10531-017-1466-x

Lubin Y. & T. Bilde 2007. The evolution of sociality in spiders. *Advances in the Study of Behavior* 37:83-145.

Madin J.S., K.D. Anderson, M.H. Andreasen, T.C. Bridge, S.D. Cairns, S.R. Connolly et al. 2016a. The Coral Trait Database, a curated database of trait information for coral species from the global oceans. *Scientific Data* 3:160017.

Madin J.S., M.O. Hoogenboom, S.R. Connolly, E.S. Darling, D.S. Falster, D. Huang et al. 2016b. A trait-based approach to advance coral reef science. *Trends in ecology and evolution* 31:419-428.

Mammola S., F. Milano, M. Vignal, J. Andrieu & M. Isaia 2019. Associations between habitat quality, body size and reproductive fitness in the alpine endemic spider Vesubia jugorum. *Global Ecology and Biogeography*

Michalko R., S. Pekár, M. Dul'a & M.H. Entling 2019a. Global patterns in the biocontrol efficacy of spiders: A meta-analysis. *Global Ecology and Biogeography* 28:1366-1378. doi:10.1111/geb.12927

Michalko R., S. Pekár, M. Dul'a & M.H. Entling 2019b. Global patterns in the biocontrol efficacy of spiders: A meta-analysis. *Global Ecology and Biogeography*

Moretti M., A.T. Dias, F. De Bello, F. Altermatt, S.L. Chown, F.M. Azcárate et al. 2017. Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. *Functional Ecology* 31:558-567.

Nentwig W., D. Gloor & C. Kropf 2015. Spider taxonomists catch data on web. *Nature* 528:479. doi:10.1038/528479a

Nentwig W, T. Blick, R. Bosmans, D. Gloor, A. Hänggi & C. Kropf 2020. Spiders of Europe. Version 1.2020. Online at https://www.araneae.nmbe.ch, accessed on 29/01/2020. https://doi.org/10.24436/1

Oliveira B.F., V.A. São-Pedro, G. Santos-Barrera, C. Penone & G.C. Costa 2017. AmphiBIO, a global database for amphibian ecological traits. *Scientific data* 4:170123.

Parr C.L., R.R. Dunn, N.J. Sanders, M.D. Weiser, M. Photakis, T.R. Bishop et al. 2017. GlobalAnts: a new database on the geography of ant traits (Hymenoptera: Formicidae). *Insect Conservation and Diversity* 10:5-20.

Pekár S. & S. Toft 2015. Trophic specialisation in a predatory group: the case of prey-specialised spiders (Araneae). *Biological Reviews* 90:744-761.

Pey B., J. Nahmani, A. Auclerc, Y. Capowiez, D. Cluzeau, J. Cortet et al. 2014. Current use of and future needs for soil invertebrate functional traits in community ecology. *Basic and Applied Ecology* 15:194-206.

Pineda S.S., P.A. Chaumeil, A. Kunert, Q. Kaas, M.W.C. Thang, L. Le et al. 2018. ArachnoServer 3.0: an online resource for automated discovery, analysis and annotation of spider toxins. *Bioinformatics* 34:1074-1076. doi:10.1093/bioinformatics/btx661

Ramírez M.J. & P. Michalik 2019. The Spider Anatomy Ontology (SPD)—A Versatile Tool to Link Anatomy with Cross-Disciplinary Data. *Diversity* 11:202.

Ratnasingham S. & P.D. Hebert 2007. BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). *Molecular ecology notes* 7:355-364.

Rusch A., K. Birkhofer, R. Bommarco, H.G. Smith & B. Ekbom 2015. Predator body sizes and habitat preferences predict predation rates in an agroecosystem. *Basic Applied Ecology* 16:250-259.

Turnbull A. 1973. Ecology of the true spiders (Araneomorphae). *Annual review of entomology* 18:305-348.

Uetz G.W., J. Halaj & A.B. Cady 1999. Guild structure of spiders in major crops. *Journal of Arachnology*:270-280.

Violle C., M.-L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel et al. 2007. Let the concept of trait be functional! 116:882-892. doi:10.1111/j.0030-1299.2007.15559.x

Wilkinson M.D., M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3

Wolff J.O., D. Wells, C.R. Reid & S.J. Blamires 2017. Clarity of objectives and working principles enhances the success of biomimetic programs. *Bioinspiration & biomimetics* 12:051001.

Wolz M., M. Klockmann, T. Schmitz, S. Pekár, D. Bonte & G. Uhl 2020. Dispersal and life-history traits in a spider with rapid range expansion. *Movement Ecology* 8:1-11.

Wong M.K., B. Guénard & O.T. Lewis 2019. Trait-based ecology of terrestrial arthropods. *Biological Reviews* 94:999-1022.

World Spider Catalog X. 2020. World Spider Catalog. Version 20.5. Natural History Museum Bern. http://wsc.nmbe.ch. Accessed 11/01/2020 2020

Wright I.J., N. Dong, V. Maire, I.C. Prentice, M. Westoby, S. Díaz et al. 2017. Global climatic drivers of leaf size. *Science* 357:917-921.

Table 1: Examples of existing databases which include data for spiders (overview compiled in January 2020).

Name	Geographic range	Trait types	Number of species included	link
Genbank	Global	DNA sequences	12537	https://www.ncbi.nlm.nih.gov/genbank/
BOLD	Global	DNA barcodes	10552	http://www.boldsystems.org/
British Spider and Harvestman Recording Schemes	Britain	Distribution, altitude, ecology, habitat	713	http://srs.britishspiders.org.uk/
Czech spider database	Czech Republic	Distribution, habitat	896	https://www.arachnology.cz/rad/18raneae-1.html
GBIF (Global Biodiversity Information Facility)	Global	Occurrence records	50708	https://www.gbif.org/
World Spider Catalog	Global	Taxonomy, species distribution	48423	https://wsc.nmbe.ch/
ArachnoServer	Global	Toxin sequence, structure, and biological activity	100	http://www.arachnoserver.org
Araneae Spiders of Europe	Europe	Distribution, morphology (body length, diagnostic illustrations), habitat	5210	https://araneae.nmbe.ch/
Global Species Database of Salticidae	Global	Distribution, morphology (diagnostic illustrations)	6990	http://www.salticidae.pl/
Jumping spiders (Arachnida: Araneae: Salticidae) of the world	Global	Distribution, morphology (diagnostic illustrations)	6149	https://www.jumping-spiders.com/
Spider Anatomy Ontology (SPD)	Global	Ontology of anatomical terms	NA	http://bioportal.bioontology.org/ontologies/SPD